



RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/809,816

Source: STIC

Date Processed by STIC: 4/2/04

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 703-308-4212; FAX: 703-308-4221

Effective 12/13/03: TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER
VERSION 4.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND
TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkr41note.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry directly to (EFFECTIVE 12/01/03):
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 10/08/03

Raw Sequence Listing Error Summary

ERROR DETECTED

SUGGESTED CORRECTION

SERIAL NUMBER: 10/809,816

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 Wrapped Nucleics
 Wrapped Aminos The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 Invalid Line Length The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 Misaligned Amino
 Numbering The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 Non-ASCII The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 Variable Length Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 PatentIn 2.0
 "bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 Skipped Sequences
 (OLD RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:
 (2) INFORMATION FOR SEQ ID NO: X: (insert SEQ ID NO where "X" is shown)
 (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: X: (insert SEQ ID NO where "X" is shown)
 This sequence is intentionally skipped

 Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 Skipped Sequences
 (NEW RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence.
 <210> sequence id number
 <400> sequence id number
 000
- 9 Use of n's or Xaa's
 (NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.
 Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
 In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10 Invalid <213>
 Response Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11 Use of <220> Sequence(s) missing the <220> "Feature" and associated numeric identifiers and responses.
 Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
 (See "Federal Register," 03/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 PatentIn 2.0
 "bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 Misuse of n/Xaa "n" can only represent a single nucleotide; "Xaa" can only represent a single amino acid



IFWO

RAW SEQUENCE LISTING

DATE: 04/02/2004

PATENT APPLICATION: US/10/809,816

TIME: 09:11:36

Input Set : A:\SEQLIST_1507.TXT

Output Set: N:\CRF4\04022004\J809816.raw

5 <110> APPLICANT: LI, Shyr-Jiann et al.
 7 <120> TITLE OF INVENTION: ISOLATED MONKEY CATHEPSIN S PROTEINS,
 8 NUCLEIC ACID MOLECULES ENCODING MONKEY CATHEPSIN S PROTEINS,
 9 AND USES THEREOF
 11 <130> FILE REFERENCE: CL001507
 C--> 13 <140> CURRENT APPLICATION NUMBER: US/10/809,816
 C--> 13 <141> CURRENT FILING DATE: 2004-03-26
 13 <160> NUMBER OF SEQ ID NOS: 11
 15 <170> SOFTWARE: FastSEQ for Windows Version 4.0

ERRORED SEQUENCES

378 <210> SEQ ID NO: 11
 379 <211> LENGTH: 24
 380 <212> TYPE: DNA
 381 <213> ORGANISM: primer
 383 <400> SEQUENCE: 11
 384 gggccctgga agcacagctg aagc
 E--> 385 7
 E--> 387 7

Does Not Comply
 Corrected Diskette Needed

(Pg. 1-3)

Invalid
Response

Mandatory <213>

24

Response has to
 be either Artificial/
 Unknown or Genus/Species

Please see item #10
 on error summary sheet.

delete

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/809,816

DATE: 04/02/2004

TIME: 09:11:38

Input Set : A:\SEQLIST_1507.TXT

Output Set: N:\CRF4\04022004\J809816.raw

L:13 M:270 C: Current Application Number differs, Replaced Current Application No
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:385 M:254 E: No. of Bases conflict, this line has no nucleotides. ✓
M:254 Repeated in SeqNo=11

10/809, 816

Page 3 of 3

<210> 7
<211> 322
<212> PRT
<213> consensus

INVALID Response
please see item #10 on
error summary sheet.

<400> 7
Met Lys Leu Val Cys Val Leu Val Cys Ser Ser Ala Val Ala Gln Leu
1 5 10 15
His Lys Asp Pro Thr Leu Asp His His Trp Leu Trp Lys Lys Thr Tyr
20 25 30
Gly Lys Gln Tyr Lys Glu Lys Asn Glu Glu Ala Val Arg Arg Leu Ile
35 40 45
Trp Glu Lys Asn Leu Lys Phe Val Met Leu His Asn Leu Glu His Ser
50 55 60
Met Gly Met His Ser Tyr Asp Leu Gly Met Asn His Leu Gly Asp Met
65 70 75 80
Thr Ser Glu Glu Val Met Ser Leu Met Ser Ser Leu Arg Val Pro Ser
85 90 95
Gln Trp Gln Arg Asn Ile Thr Tyr Lys Ser Asn Asn Gln Leu Pro Asp
100 105 110
Ser Val Asp Trp Arg Glu Lys Gly Cys Val Thr Glu Val Lys Tyr Gln
115 120 125
Gly Ser Cys Gly Ala Cys Trp Ala Phe Ser Ala Val Gly Ala Leu Glu
130 135 140
Ala Gln Leu Lys Leu Lys Thr Gly Lys Leu Val Ser Leu Ser Ala Gln
145 150 155 160
Asn Leu Val Asp Cys Ser Thr Glu Lys Tyr Gly Asn Lys Gly Cys Asn
165 170 175
Gly Gly Phe Met Thr Ala Phe Gln Tyr Ile Ile Asp Asn Gly Ile Asp
180 185 190
Ser Asp Ala Ser Tyr Pro Tyr Lys Ala Met Asp Gln Lys Cys Gln Tyr
195 200 205
Asp Ser Lys Tyr Arg Ala Ala Thr Cys Ser Lys Tyr Thr Glu Leu Pro
210 215 220
Tyr Gly Arg Glu Asp Val Leu Lys Glu Ala Val Ala Asn Lys Gly Pro
225 230 235 240
Val Ser Val Gly Val Asp Ala Ser His Pro Ser Phe Phe Leu Tyr Arg
245 250 255
Ser Gly Val Tyr Tyr Glu Pro Ser Cys Thr Gln Asn Val Asn His Gly
260 265 270
Val Leu Val Val Gly Tyr Gly Leu Asn Gly Lys Glu Tyr Trp Leu Val
275 280 285
Lys Asn Ser Trp Gly Asn Phe Gly Glu Gln Gly Tyr Ile Arg Met Ala
290 295 300
Arg Asn Lys Gly Asn His Cys Gly Ile Ala Ser Tyr Pro Ser Tyr Pro
305 310 315 320
Glu Ile

<210> 8
<211> 31
<212> DNA
<213> primer

SAME error

<400> 8
ccggaattct tgcataaaga tcccaccctg g

31
The type of errors shown exist throughout
the Sequence Listing. Please check subsequent
sequences for similar errors.